

100 kb

1 TTGGTAGCAACGGAAACCGCGGGCGCGGTTTCGGCCCGGCTCCCGGGCGGCTCCTTGCTC
61 TCGGCGGGCCTCCCCGCCCCCTTCGTGTCGTCTTCTCCCCCTCACCAGCCCCGGGCGCCC
121 CTCCGGCCGCGCCAACCCCGGCTCCCCGCTCGCGCCCCGTGCTCCCCGCGCGGTTCGG
181 GCGTCTCCTTGGCGCGCCCCGGCTCCCGGCTGTCCCCGCCCCGGCGTGCGAACCAGGTGTATG
241 SCA2-A → OGCCCCCTACCATGTCTGCTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGC
301 AGCAACCAACAGCAGCAGCAGCAGCAGCAACAGCCCGCCGCCCCGGCTGCCAATGTCCGCA ← SCA2-B
361 AGCCCGGGCGGCAGCGGCCCTTCTAGCGTCGCCCCGCCGCCCGCGCCTTCGCCGTCTCTCTCT
421 CGGTCTCTCTGTCCTCGGCCACGGCTCCCTCCTCGGTGCTCGCGGCCACCTCCGGCGCCG
481 GGAAGCCCCGGCCTGGGCAG GTGGGTGTGCGGCACCCC

FIGURE 2

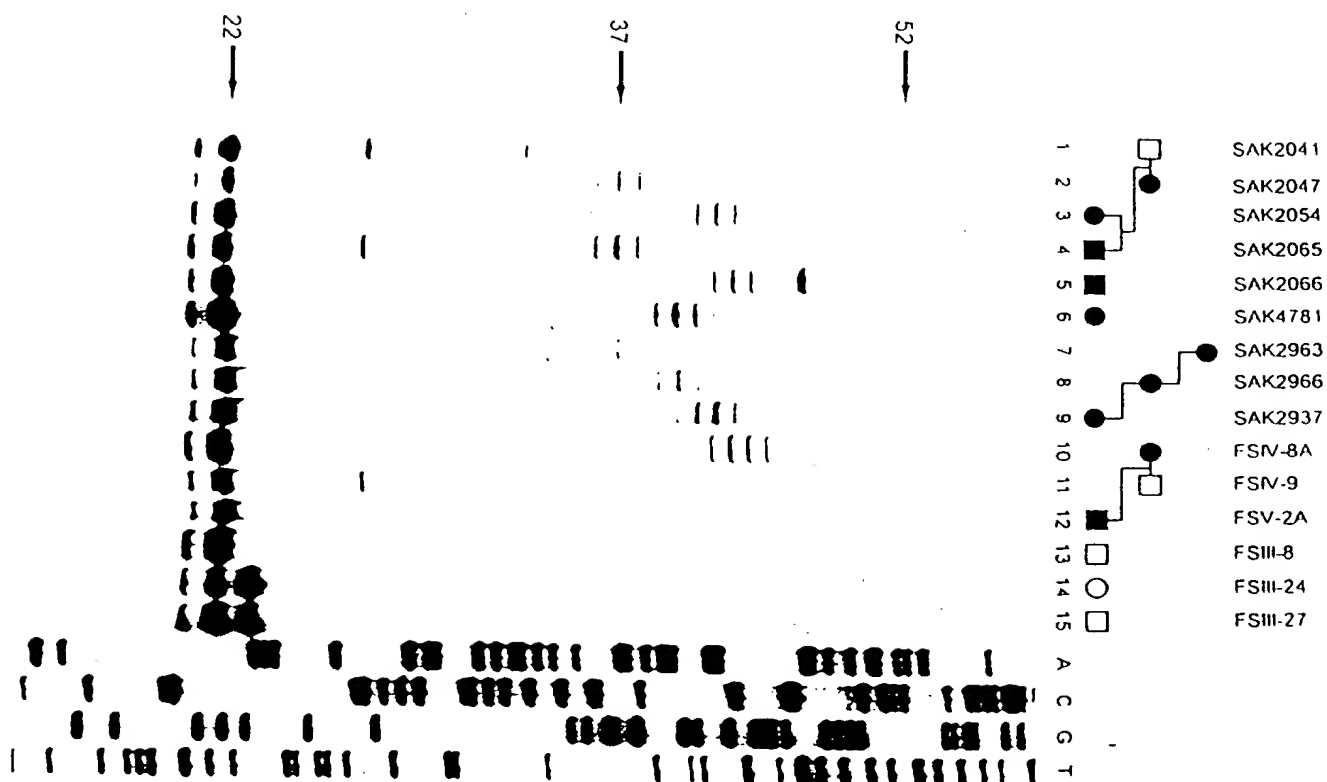


FIGURE 3

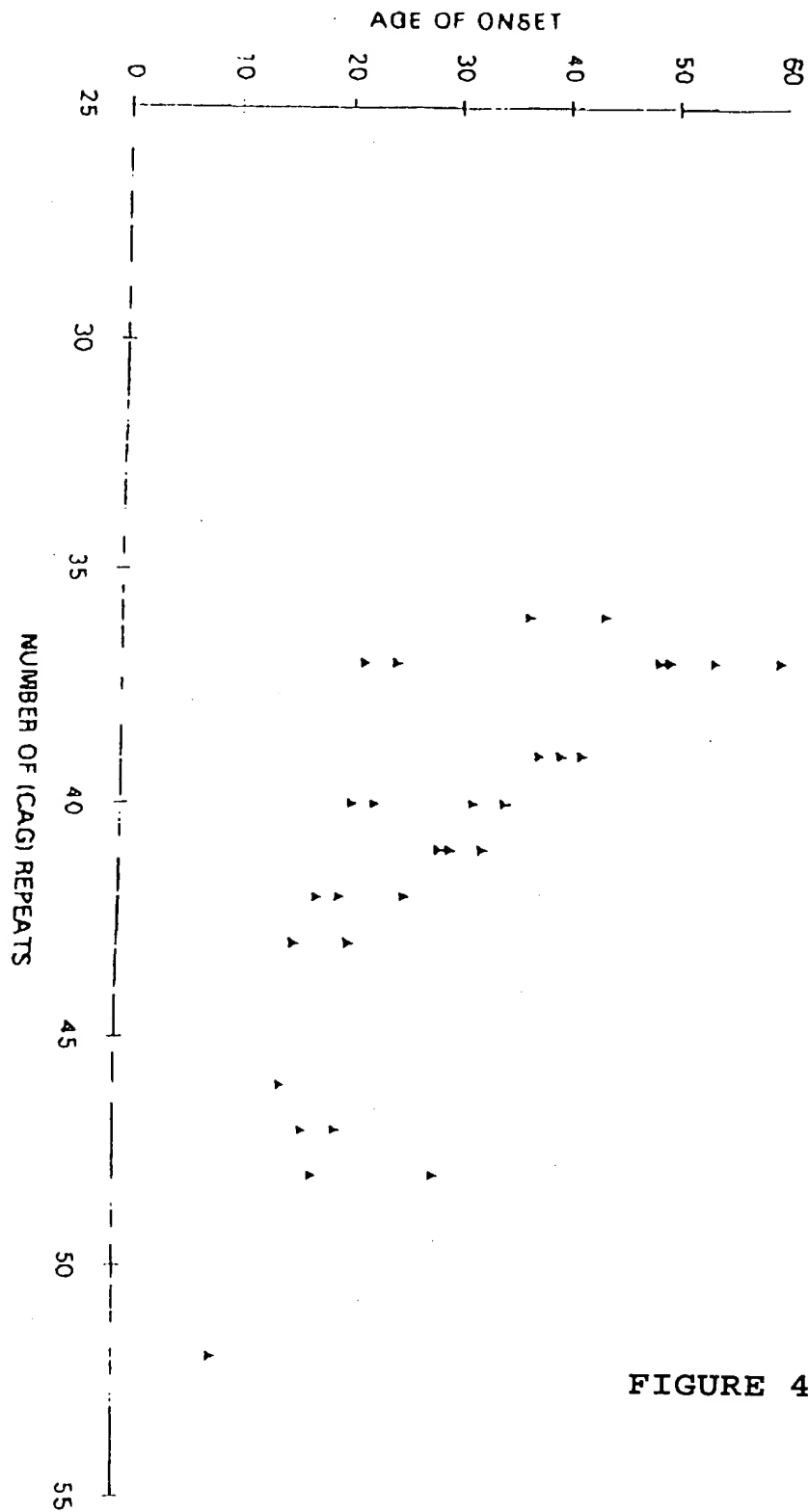


FIGURE 4

F1.1-F1.7 5' — AAA 3'

F2.1-F2.8 5' — AAA 3'

S1 5' — 3'

S2 5' — 3'

300 bp

FIGURE 5

1	ACCCCCGAGAAAGCAACCCAGCGCGCCGCCGCTCCTCACGTGTCCCTCCCGGCCCGGG	60
61	GCCACCTCACGTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120
121	CACCTCCGCTCCCACCCGGCGCCTCGGCGCGCCCCGCCCTCCGATGCGCTCAGCGGCCGCA	180
1	M R S A A A	6
181	GCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCCGCAGCCAGGTGGCCC	240
7	A P R S P A V A T E S R R F A A A R W P	26
241	GGGTGGCGCTCGCTCCAGCGGCCGCGCGGGCGGAGCGGGCGGGCGGGTGGCGCGGCC	300
27	G W R S L Q R P A R R S G R G G G A A	46
301	CCGGGACCGTATCCCTCCGCCGCCCTCCCCCGCCCGGCCCGGCCCTCCCTCCCGG	360
47	P G P Y P S A A P P P P G P G P P P S R	66
361	CAGAGCTCGCCTCCCTCCGCCTCAGACTGTTTTGGTAGCAACGGCAACGGCGGGCGCGC	420
67	Q S S P P S A S D C F G S N G G A	86
421	TTTCGGCCCCGGCTCCCGGCGGCTCCTTGGTCTCGGCGGGCCTCCCCGCCCTTCGTCGTC	480
87	F R P G S R R L L G L G G P P R P F V V	106
481	GTCCTTCTCCCCCTCGCCAGCCCGGGCGCCCTCCGGCCGCGCCAACCCGCGCCTCCCCG	540
107	V L L P L A S P G A P P A A P T R A S P	126
541	CTCGGCGCCCGTGCCTCCCGCCGCGTTCCGGCGTCTCCTTGGCGCGCCCGGCTCCCGGC	600
127	L G A R A S P P R S G V S L A R P A P G	146
	SCA2-A	
601	TGTCCCCGCCCCGGCGTGCGAGCCGGTGTATGGGCCCTCACCATGTCGCTGAAGCCCCAG	660
147	C P R P A C E P V Y G P L T M S L K P Q	166
661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG	720
167	Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	186
	SCA2-B	
721	CAGCCGCCGCCCCGGCTGCCAATGTCCGCAAGCCCCGGCGGCAGCGGCCTTCTAGCGTCG	780
187	Q P P P A A A N V R K P G G S G L L A S	206
781	CCCGCCGCCGCGCCTTCGCCGTCCTCGTCTCCTCGTCTCCTCGTCTCGGCCACGGCTCCC	840
207	P A A A P S P S S S S V S S S S A T A P	226
841	TCCTCGGTGGTCCGCGCGACCTCCGGCGGGGAGGCCCCGGCCTGGGCAGAGGTGAAAC	900
227	S S V A A T S G G G R P G L G R G R N	246
901	AGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAATAGAGG	960
247	S N K G L P Q S T I S F D G I Y A N M R	266
961	ATGGTTTCATATACTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA	1020
267	M V H I L T S V V G S K C E V Q V K N G	286
	SCA2-14B	
1021	GGTATATATGAAGGAGTTTTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080
287	G I Y E G V F K T Y S P K C D L V L D A	306
1081	GCACATGAGAAAAGTACAGAATCCAGTTCGGGGCCGAAACGTGAAGAAATAATGGAGAGT	1140
307	A H E K S T E S S S G P K R E E I M E S	326
1141	ATTTTGTTCAAATGTTTCAGACTTTGTTGTGGTACAGTTTAAAGATATGGACTCCAGTTAT	1200
327	I L F K C S D F V V V Q F K D M D S S Y	346
1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAAGTGAATGGCGAACACAAA	1260
347	A K R D A F T D S A I S A K V N G E H K	366
1261	GAGAAGGACCTGGAGCCCTGGGATGCAGGTGAACCTCACAGCCAATGAGGAACCTTGAGGCT	1320
367	E K D L E P W D A G E L T A N E E L E A	386
1321	TTGGAAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA	1380
387	L E N D V S N G W D P N D M F R Y N E	406
1381	AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCCTTAGAA	1440
407	N Y G V V S T Y D S S L S S Y T V P L E	426
1441	AGAGATAACTCAGAAGAATTTTTTAAACGGGAAGCAAGGGCAAACCAAGTTAGCAGAAGAA	1500
427	R D N S E E F L K R E A R A N Q L A E E	446

FIGURE 6A

1501	ATTGAGTCAAGTGCCCAAGTACAAAAGCTCGAGTGGCCCTGGAAAATGATGATAGGAGTGAG	1560
447	I E S S A Q Y K A R V A L E N D D R S E	466
1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC	1620
467	E E K Y T A V Q R N S S E R E G H S I N	486
1621	ACTAGGGAAAATAAATATATTCCTCCTGGACAAAGAAATAGAGAAGTCATATCCTGGGGA	1680
487	T R E N K Y I P P G Q R N R E V I S W G	506
1681	AGTGGGAGACAGAATTCACCGCGTATGGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA	1740
507	S G R Q N S P R M G Q P G S G S M P S R	526
1741	TCCACTTCTCACACTTCAGATTTCAACCCGAATTCTGGTTCAGACCAAAGAGTAGTTAAT	1800
527	S T S H T S D F N P N S G S D Q R V V N	546
1801	GGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCCTTCTCTCGCCACCTTCTCGCTAC	1860
547	G G V P W P S P C P S P S S R P P S R Y	566
1861	CAGTCAGGTCCCAACTCTCTTCCACCTCGGGCAGCCACCCCTACACGGCCGCCCTCCAGG	1920
567	Q S G P N S L P P R A A T P T R P P S R	586
1921	CCCCCTCGCGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGCTCCT	1980
587	P P S R P S R P P S H P S A H G S P A P	606
1981	GTCTCTACTATGCCTAAACGCATGTCTTCAGAAGGGCCTCCAAGGATGTCCCCAAAGGCC	2040
607	V S T M P K R M S S E G G P P R M S P K A	626
2041	CAGCGACATCCTCGAAATCACAGAGTTTCTGCTGGGAGGGGTTCCATATCCAGTGGCCTA	2100
627	Q R H P R N H R V S A G R G S I S S G L	646
2101	GAATTTGTATCCCAACAACCCACCCAGTGAAGCAGCTACTCCTCCAGTAGCAAGGACCAGT	2160
647	E F V S H N P P S E A A T P P V A R T S	666
2161	CCCTCGGGGGGAACGTGGTCATCAGTGGTCAGTGGGGTTCCAAGATTATCCCCTAAACT	2220
667	P S G G T W S S V V S G V P R L S P K T	686
2221	CATAGACCCAGGTCTCCCAGACAGAACAGTATTGGAATACCCCAAGTGGGCCAGTTCTT	2280
687	H R P R S P R Q N S I G N T P S G P V L	706
2281	GCTTCTCCCCAAGCTGGTATTATTCCAAGTGAAGCTGTTGCCATGCCTATTCCAGCTGCA	2340
707	A S P Q A G I I P T E A V A M P I P A A	726
2341	TCTCTACGCCTGCTAGTCTGCATCGAACAGAGCTGTTACCCCTTCTAGTGAGGCTAAA	2400
727	S P T P A S P A S N R A V T P S S E A K	746
2401	GATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCCTGCAGGGAATAAAGAAAATATTTAAA	2460
747	D S R L Q D Q R Q N S P A G N K E N I K	766
2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAAACAAAGGTATATCACCAGTTGTT	2520
767	P N E T S P S F S K A E N K G I S P V V	786
2521	TCTGAACATAGAAAACAGATTGATGATTAAAGAAATTTAAGAATGATTTTAGGTTACAG	2580
787	S E H R K Q I D D L K K F K N D F R L Q	806
2581	CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAATAGAGAGGGAGAAAAA	2640
807	P S S T S E S M D Q L L N K N R E G E K	826
2641	TCAAGAGATTTGATCAAAGACAAAATGAACCAAGTGCTAAGGATTCTTTCATTGAAAAT	2700
827	S R D L I K D K I E P S A K D S F I E N	846
2701	AGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCTTCA	2760
847	S S S N C T S G S S K P N S P S I S P S	866
2761	ATACTTAGTAACACGGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTACAGCT	2820
867	I L S N T E H K R G P E V T S Q G V Q T	886
2821	TCCAGCCCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAGAAGAAAGACGCAGCTGAG	2880
887	S S P A C K Q E K D D K E E K K D A A E	906
2881	CAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTCAACCCACGTTCTTCTCT	2940
907	Q V R K S T L N P N A K E F N P R S F S	926
2941	CAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCAAGCACACCTAGCCCATCT	3000
927	Q P K P S T T P T S P R P Q A Q P S P S	946
3001	ATGGTGGGTCAACAGCCAACTCCAGTTTATCTCAGCCTGTTTGTGTTGCACCAAT	3060
947	M V G H Q Q P T P V Y T Q P V C F A P N	966
3061	ATGATGTATCCAGTCCCAGTGAGCCAGGCGTGCAACCTTTATACCCAATACCTATGACG	3120
967	M M Y P V P V S P G V Q P L Y P I P M T	986

FIGURE 6B

3121	CCCATGCCAGTGAATCAAGCCAAGACATATAGAGCAGTACCAAATATGCCCCAACAGCGG	3180
987	P M P V N Q A K T Y R A V P N M P Q Q R	1006
3181	CAAGACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCAGCAGCGGGCCACCGATT	3240
1007	Q D Q H H Q S A M M H P A S A A G P P I	1026
3241	GCAGCCACCCCAACAGCTTACTCCACGCAATATGTTGCCTACAGTCCTCAGCAGTTCCCA	3300
1027	A A T P P A Y S T Q Y V A Y S P Q Q F P	1046
3301	AATCAGCCCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAGCATCCTCATGTCTATAGT	3360
1047	N Q P L V Q H V P H Y Q S Q H P H V Y S	1066
3361	CCTGTAATACAGGGTAATGCTAGAAATGATGGCACCACCAACACACGCCCAGCCTGGTTTA	3420
1067	P V I Q G N A R M M A P P T H A Q P G L	1086
3421	GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGCATGT	3480
1087	V S S S A T Q Y G A H E Q T H A M Y A C	1106
3481	CCCAAATTACCATAACAAGGAGACAAGCCCTTCTTTCTACTTTGCCATTTCCACGGGC	3540
1107	P K L P Y N K E T S P S F Y F A I S T G	1126
3541	TCCCTTGCTCAGCAGTATGCGCACCCCTAACGCTACCCTGCACCCACATACTCCACACCCT	3600
1127	S L A Q Q Y A H P N A T L H P H T P H P	1146
3601	CAGCCTTCAGCTACCCCCACTGGACAGCAGCAAAGCCAACATGGTGGAAGTCATCCTGCA	3660
1147	Q P S A T P T G Q Q Q S Q H G G S H P A	1166
3661	CCCAGTCCTGTTTCAGCACCATCAGCACCAGGCCGCCAGGCTCTCCATCTGGCCAGTCCA	3720
1167	P S P V Q H H Q H Q A A Q A L H L A S P	1186
3721	CAGCAGCAGTCAGCCATTTACCACGCGGGGCTTGCGCCAACCTCCACCCTCCATGACACCT	3780
1187	Q Q Q S A I Y H A G L A P T P P S M T P	1206
3781	GCCTCCAACACGCAGTCGCCACAGAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG	3840
1207	A S N T Q S P Q N S F P A A Q Q T V F T	1226
3841	ATCCATCCTTCTCAGTTTCAGCCGGCGTATACCAACCCACCCACATGGCCCCACGTACCT	3900
1227	I H P S H V Q P A Y T N P P H M A H V P	1246
3901	CAGGCTCATGTACAGTCAGGAATGTTCTTCTCATCCAACCTGCCCATGCGCCAATGATG	3960
1247	Q A H V Q S G M V P S H P T A H A P M M	1266
3961	CTAATGACGACACAGCCACCCGGCGGTCCCCAGGCCGCCCTCGCTCAAAGTGCACTACAG	4020
1267	L M T T Q P P G G P Q A A L A Q S A L Q	1286
4021	CCCATTCCAGTCTCGACAACAGCGCATTTCCTTATATGACGCACCCTTCAGTACAAGCC	4080
1287	P I P V S T T A H F P Y M T H P S V Q A	1306
4081	CACCACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCCTCCTC	4140
1307	H H Q Q Q L *	1326
4141	CCTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACCTAGAAATTTTATTTTGTTTT	4200
4201	TAAAATATATATGTTGATTTCTTGTAACATCCAATAGGAATGCTAACAGTTCACTTGCAG	4260
4261	TGGAAGATACTTGGACCGAGTAGAGGCATTTAGGAACCTGGGGGCTATTCCATAATTCCA	4320
4321	TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCCGAAACTGGAAGTTATTT	4380
4381	ATTTTTTAATAACCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAGTAACAAGAGT	4440
4441	GATTCTTGCTGCTATTACTGCTAAAAAAAAAAAAAAAAAAAAA 4481	

FIGURE 6C

	1				50
Ataxin-2	VYGPLTMSLK	PQQQQQQQQQ	QQQQQQQQQQ	QQQPPPAAN	VRKPGGSGLL
Mouse Ataxin-2	HEGPLTMSLK	PQPQ.....	PPAPAT	GRKPGG.GLL
A2RPLA	PQPPPPQHQ	ER.....
Consensus	-----L-	PQ-----	-----	-----	-----
	51				100
Ataxin-2	ASPAAAPSPS	SSSVSSSSAT	APSSVVA...	ATSGGGRPGL	GRGRNSNKGL
Mouse Ataxin-2	SSPGAAP.AS	AAVTSASVVP	APAAPVASSS	AAAGGGRPGL	GRGRNSSKGL
A2RP	..PGAAAIGS	A.....RGQSTGKGP
Consensus	--P-AA---S	-----	-----	-----	-RG----KG-
	101				150
Ataxin-2	PQSTISFDGI	YANMRMVHIL	TSVVGSKCEV	QVKNGGIYEG	VFKTYSPKCD
Mouse Ataxin-2	PQPTISFDGI	YANVRMVHIL	TSVVGSKCEV	QVKNGGIYEG	VFKTYSPKCD
A2RP	PQSPV.FEGV	YNNRMLHFL	TAVVGSTCDV	KVKNGTTYEG	IFKTLSSKFE
Consensus	PQ----F-G-	Y-N-RM-H-L	T-VVGS-C-V	-VKNG--YEG	-FKT-S-K--
	151				200
Ataxin-2	LVLDAAEKS	TESSSGPKRE	EIMESILFKC	SDFVVVQFKD	MDSSYAKRDA
Mouse Ataxin-2	LVLDAAEKS	TESSSGPKRE	EIMESVLFKC	SDFVVVQFKD	TDSSYARRDA
A2RP	LAVDAVHRKA	SEPAGGPRRE	DIVDTMVFKP	SDVMLVHFRN	VDFNYATKDK
Consensus	L--DA-H-K-	-E---GP-RE	-I-----FK-	SD---V-F--	-D--YA--D-
	201				250
Ataxin-2	FTDSAIS..A	KVNGEHKEKD	LEPWDAGELT	ANEELEALEN	DVSNGWDPND
Mouse Ataxin-2	FTDSALS..A	KVNGEHKEKD	LEPWDAGELT	ASEELE.LEN	DVSNGWDPND
A2RP	FTDSAIAMNS	KVNGEHKEKV	LQRWEGGD.S	NSDDYD.LES	DMSNGWDPNE
Consensus	FTDSA-----	KVNGEHKEK-	L--W--G---	-----LE-	D-SNGWDPN-
	251				300
Ataxin-2	MFRYNEENYG	VVSTYDSSLS	SYTVPLERDN	SEEFLKREAR	ANQLAEEIES
Mouse Ataxin-2	MFRYNEENYG	VVSTYDSSLS	SYTVPLERDN	SEEFLKREAR	ANQLAEEIES
A2RP	MFKFNEENYG	VKTTYDSSLS	SYTVPLEKDN	SEEFROREL	AAQLAREIES
Consensus	MF--NEENYG	V--TYDSSLS	SYTVPLE-DN	SEEF--RE-R	A-QLA-EIES
	301				350
Ataxin-2	SAQYKARVAL	ENDD.RSEEE	KYTAVQRNCS	EREGHSINTR	ENKYIPPGQR
Mouse Ataxin-2	SAQYKARVAL	ENDD.RSEEE	KYTAVQRNCS	DREGHGPNTR	DNKYIPPGQR
A2RP	SPQYRLRIAM	ENDDGRTEEE	KHSAVQRQGS	GRESPLASR	EGKYIP....
Consensus	S-QY--R-A-	ENDD-R-EEE	K--AVQR--S	-RE-----R	--KYIP----
	351				
Ataxin-2	NR				
Mouse Ataxin-2	NR				
A2RP	..				
Consensus	--				

FIGURE 7